Markov Processes Relat. Fields **20**, 545–562 (2014)

Markov Processes and Related Fields

From Level 2.5 to Level 2 Large Deviations for Continuous Time Markov Chains

L. Bertini[1], A. Faggionato[1] and D. Gabrielli[2]

[1] Dipartimento di Matematica, Università di Roma 'La Sapienza' P.le Aldo Moro 2, 00185 Roma, Italy. E-mail: bertini@mat.uniroma1.it, faggiona@mat.uniroma1.it
[2] Dipartimento di Matematica, Università dell'Aquila Via Vetoio, 67100 Coppito, L'Aquila, Italy. E-mail: gabriell@univaq.it

Received June 11, 2013

Abstract. We provide a new variational characterization of the rate function for the Donsker – Varadhan large deviations principle (LDP) of the empirical measure of a continuous time Markov chain on a countable (finite or infinite) state space. This is obtained by contraction from the joint LDP for the empirical measure and the empirical flow proved in [2].

Keywords: Markov chain, Large deviations principle, Contraction principle, Empirical flow, Empirical measure, Fenchel – Rockafellar Theorem

AMS Subject Classification: Primary 60F10, 60J27; Secondary 82C05

1. Introduction

We consider a continuous time Markov chain $(\xi_t)_{t\geq 0}$ on a countable (finite or infinite) state space V. Following [9] the dynamics is defined knowing the *jump rates* $r(x,y)$, $x \neq y$ in V, under the assumption that $r(x) := \sum_{y\in V} r(x,y) < +\infty$ for all $x \in V$. Then, at each site x the system waits an exponential time of parameter $r(x)$ afterwards it jumps to a state y with probability $r(x,y)/r(x)$. We assume that a.s. for any fixed initial state explosion does not occur, hence the Markov chain is defined in V for all times $t \in \mathbb{R}_+$ and we do not need to introduce any coffin state. We denote by $\mathbb{P}_x$ the law on the Skorohod space $D(\mathbb{R}_+; V)$ of the Markov chain starting at x.

In what follows we restrict to irreducible Markov chains such that there exists a unique invariant probability measure, which we denote by π. As in [9], by invariant probability measure π we mean a probability measure on V such that

$$\sum_{y\in V}\pi(x)\,r(x,y)=\sum_{y\in V}\pi(y)\,r(y,x)\qquad \forall\, x\in V \tag{1.1}$$

where we understand $r(x,x)=0$. We stress that the existence of π is guaranteed if V is finite, while in general uniqueness is automatic if π exists.

A fundamental result in the theory of large deviations is given by the Donsker-Varadhan Large Deviations Principle (LDP) of the empirical measure of Markov processes. We recall its statement referring to the above Markov chain ξ. Denote by $\mathcal{P}(V)$ the space of probability measures on V endowed of the weak topology. Given $T>0$ the *empirical measure* $\mu_T\colon D(\mathbb{R}_+;V)\to\mathcal{P}(V)$ is defined by

$$\mu_T\,(X)=\frac{1}{T}\int_0^T dt\,\delta_{X_t} \tag{1.2}$$

where δ_y denotes the pointmass at y. Given $x\in V$, the ergodic theorem [9] implies that the empirical measure μ_T converges $\mathbb{P}_x$ a.s. to π as $T\to\infty$. In particular, the family of probabilities $\{\mathbb{P}_x\circ\mu_T^{-1}\}_{T>0}$ on $\mathcal{P}(V)$ converges to δ_π. In [6] the large deviations from the above limit theorem have been studied by Donsker and Varadhan. Under suitable hypotheses (see Remark 1.2 below) they proved that as $T\to+\infty$ the family of probability measures $\{\mathbb{P}_x\circ\mu_T^{-1}\}_{T>0}$ on $\mathcal{P}(V)$ satisfies a LDP with good rate function $\mathcal{I}$ such that

$$\mathcal{I}(\mu)=\sup_h\left\{-\langle\mu,h^{-1}Lh\rangle\right\} \tag{1.3}$$

as h varies among the strictly positive functions in the domain of the infinitesimal generator L (in general, $<\mu,f>:=\sum_{x\in V}\mu(x)f(x)$).

The above result has been derived in [6]–(I) from an analogous result for discrete time Markov chains by an approximation argument in the case of V finite. The extension to V infinite has been achieved in [6]–(III), while in [6]–(IV) the LDP for the empirical measure is obtained by contraction from the LDP for the empirical process.

Our aim in this note is to give an alternative proof of the LDP for the empirical measure by contraction from the joint LDP for the empirical measure and flow recently proved in [2]. As a consequence, we derive a new representation of the rate function for the empirical measure. The classical Donsker – Varadhan representation is given as a supremum of an explicit functional over a suitable class of functions (cf. equation (1.3)). Our representation is given as an infimum of an explicit functional over a suitable class of flows. In [6]–(IV) there is also a representation in term of an infimum but the functional to be minimized is not really explicit. Combining the two representations, the classic one and ours, one gets easily upper and lower bounds on the rate function by choosing suitable test flows and functions, respectively. The equivalence of the two representations follows at once since they both correspond to the rate function of the same random object, the empirical measure. We give an alternative and direct proof of this fact, independent from the LDP proved by Donsker and Varadhan in [6]. This proof exploits several discrete geometric features concerning the graph underlying the Markov chain and the associated divergence-free flows, which can be interesting by themselves. We show also that in the finite dimensional case the equivalence of the two representations of the rate function is an instance of the Fenchel−Rockafellar duality. The infinite dimensional case is more subtle, cannot be achieved by adapting the finite-dimensional proof and requires new arguments.

We recall the joint LDP for the empirical measure and flow in [2] and fix some notation. We denote by E the set of ordered edges in V with positive transition rate, namely $E := \{(y, z) \in V \times V : y \neq z \text{ and } r(y, z) > 0\}$. Then for each $T > 0$ we define the *empirical flow* as the map $Q_T : D(\mathbb{R}_+; V) \to [0, +\infty]^E$ given by

$$Q_T(y, z)\,(X) := \frac{1}{T} \sum_{0 \leq t \leq T} \delta_y(X_{t^-})\delta_z(X_t) \qquad (y, z) \in E\,. \tag{1.4}$$

Namely, $TQ_T(y, z)$ is $\mathbb{P}_x$ a.s. the number of jumps from y to z in the time interval $[0, T]$ of the Markov chain ξ starting at x. As discussed in [2], $Q_T(y, z)$ converges to $\pi(y)r(y, z)$ at $T \to \infty$ $\mathbb{P}_x$ a.s.

Elements in $[0, +\infty]^E$ are called *flows*. We denote by $L^1_+(E)$ the subset of *summable flows*, i.e. of flows Q such that $\|Q\|_1 := \sum_{(y,z)\in E} Q(y, z) < +\infty$. Given a summable flow $Q \in L^1_+(E)$ its *divergence* $\operatorname{div} Q : V \to \mathbb{R}$ is defined as

$$\operatorname{div} Q\,(y) = \sum_{z:\,(y,z)\in E} Q(y, z) - \sum_{z:\,(z,y)\in E} Q(z, y), \qquad y \in V. \tag{1.5}$$

Observe that the divergence maps $L^1_+(E)$ into $L^1(V)$.

To each probability $\mu \in \mathcal{P}(V)$ we associate the flow $Q^\mu \in \mathbb{R}^E_+$ defined by

$$Q^\mu(y, z) := \mu(y)\, r(y, z) \qquad (y, z) \in E. \tag{1.6}$$

Note that $Q^\mu \in L^1_+(E)$ if and only if $\langle \mu, r\rangle < +\infty$. Moreover, in this case, by (1.1) Q^μ has vanishing divergence if only if μ is invariant for the Markov chain ξ, i.e. $\mu = \pi$.

We endow $L^1_+(E)$ of the bounded weak* topology. As discussed in [2] this topology is the most suited for studying large deviations of the empirical flow. For completeness we recall its definition although it will never be used below (see [8] for a detailed treatment). A subset $W \subset L^1_+(E)$ is open if and only if for each $\ell > 0$ the set $\{Q \in W : \|Q\|_1 < \ell\}$ is open in the ball $\{Q \in L^1_+(E) : \|Q\|_1 < \ell\}$

endowed of the weak* topology inherited from $L^1(E)$. When E is finite, the bounded weak* topology coincides with the L^1-topology.

We can now recall the LDP proved in [2]. We start from the assumptions. To this aim, given $f\colon V \to \mathbb{R}$ such that $\sum_{y\in V} r(x,y)\,|f(y)| < +\infty$ for each $x \in V$, we denote by $Lf\colon V \to \mathbb{R}$ the function defined by

$$Lf\,(x) := \sum_{y\in V} r(x,y)\big[f(y) - f(x)\big], \qquad x \in V. \tag{1.7}$$

Definition 1.1. Given $\sigma \in \mathbb{R}_+$ we say that Condition $C(\sigma)$ holds if there exists a sequence of functions $u_n\colon V \to (0,+\infty)$ satisfying the following requirements:

(i) For each $x \in V$ and $n \in \mathbb{N}$ it holds $\sum_{y\in V} r(x,y)u_n(y) < +\infty$.

(ii) The sequence u_n is uniformly bounded from below. Namely, there exists $c > 0$ such that $u_n(x) \geq c$ for any $x \in V$ and $n \in \mathbb{N}$.

(iii) The sequence u_n is uniformly bounded from above on compacts. Namely, for each $x \in V$ there exists a constant C_x such that for any $n \in \mathbb{N}$ it holds $u_n(x) \leq C_x$.

(iv) Set $v_n := -Lu_n/u_n$. The sequence $v_n\colon V \to \mathbb{R}$ converges pointwise to some $v\colon V \to \mathbb{R}$.

(v) The function v has compact level sets. Namely, for each $\ell \in \mathbb{R}$ the level set $\big\{x \in V\,:\, v(x) \leq \ell\big\}$ is finite.

(vi) There exists a positive constant C such that $v \geq \sigma\, r - C$.

Several examples of Markov chains satisfying condition $C(\sigma)$ with $\sigma > 0$ are discussed in [2, 3].

Let $\Phi\colon \mathbb{R}_+ \times \mathbb{R}_+ \to [0,+\infty]$ be the function defined by

$$\Phi(q,p) := \begin{cases} q\log\dfrac{q}{p} - (q-p) & \text{if } q,p \in (0,+\infty), \\ p & \text{if } q = 0,\ p \in [0,+\infty), \\ +\infty & \text{if } p = 0 \text{ and } q \in (0,+\infty). \end{cases} \tag{1.8}$$

For $p > 0$, $\Phi(\cdot,p)$ is a nonnegative convex function and is zero only at $q = p$. Indeed, it is the rate function for the LDP of the sequence N_T/T as $T \to +\infty$, $(N_t)_{t\in\mathbb{R}_+}$ being a Poisson process with parameter p.

Finally, we let $I\colon \mathcal{P}(V) \times L^1_+(E) \to [0,+\infty]$ be the functional defined by

$$I(\mu,Q) := \begin{cases} \displaystyle\sum_{(y,z)\in E} \Phi\big(Q(y,z), Q^\mu(y,z)\big) & \text{if } \operatorname{div} Q = 0\,,\ \langle \mu, r\rangle < +\infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{1.9}$$

Remark 1.1. As proved in [2, Appendix B] the above condition $\langle \mu, r\rangle < +\infty$ can be removed, since the series in (1.9) diverges if $\langle \mu, r\rangle = +\infty$.

Remark 1.2. Condition $C(0)$ (i.e. $C(\sigma)$ with $\sigma = 0$) with (i) replaced by the fact that u_n belongs to the domain of the infinitesimal generator, and with Lu_n defined as the infinitesimal generator applied to u_n, is the condition under which the large deviation of the empirical measure is derived in [6]–(IV).

Theorem 1.1 (Bertini, Faggionato, Gabrielli, [2]). *Assume Condition $C(\sigma)$ to hold with $\sigma > 0$. (Alternatively, assume the hypercontractivity Condition 2.3 in* [2]*). Then as $T \to +\infty$ the family of probability measures $\{\mathbb{P}_x \circ (\mu_T, Q_T)^{-1}\}$ on $\mathcal{P}(V) \times L^1_+(E)$ satisfies a LDP with good and convex rate function I. Namely, for each closed set $\mathcal{C} \subset \mathcal{P}(V) \times L^1_+(E)$, and each open set $\mathcal{A} \subset \mathcal{P}(V) \times L^1_+(E)$, it holds for each $x \in V$*

$$\varlimsup_{T\to+\infty} \frac{1}{T} \log \mathbb{P}_x\Big((\mu_T, Q_T) \in \mathcal{C}\Big) \leq - \inf_{(\mu,Q)\in\mathcal{C}} I(\mu, Q), \tag{1.10}$$

$$\varliminf_{T\to+\infty} \frac{1}{T} \log \mathbb{P}_x\Big((\mu_T, Q_T) \in \mathcal{A}\Big) \geq - \inf_{(\mu,Q)\in\mathcal{A}} I(\mu, Q). \tag{1.11}$$

Remark 1.3. Condition $C(\sigma)$ with $\sigma > 0$ (or alternatively the hypercontractivity Condition 2.3 in [2]) implies that $\langle \pi, r\rangle < +\infty$ (see Lemma 3.9 in [2]).

We can finally state our new results.

Theorem 1.2. *Assume that Condition $C(\sigma)$ holds with $\sigma > 0$ (alternatively, assume the hypercontractivity Condition 2.3 in [2]). Then as $T \to +\infty$ the family of probability measures $\{\mathbb{P}_x \circ \mu_T^{-1}\}$ on $\mathcal{P}(V)$ satisfies a LDP with good rate function $\mathcal{I}$ such that*

$$\mathcal{I}(\mu) = \inf\big\{I(\mu, Q) \,:\, Q \in L^1_+(E)\big\}\,. \tag{1.12}$$

$\mathcal{I}(\mu) < +\infty$ if and only if $\langle \mu, r\rangle < +\infty$, in this case the above infimum is indeed attained at a unique flow $Q^* \in L^1_+(E)$. Moreover the following alternative variational characterization holds

$$\mathcal{I}(\mu) = \begin{cases} \sup\big\{ - \langle \mu, e^{-g} L e^{g}\rangle \,:\, g \in L^\infty(V)\big\} & \textit{if } \langle \mu, r\rangle < +\infty\,, \\ +\infty & \textit{otherwise}\,. \end{cases} \tag{1.13}$$

Since the projection map $\mathcal{P}(V) \times L^1_+(E) \ni (\mu, Q) \to \mu \in \mathcal{P}(V)$ is trivially continuous, due to the contraction principle the first part of the theorem up to (1.12) follows from Theorem 1.1. Since $I(\mu, Q) = +\infty$ if $\langle \mu, r\rangle = +\infty$ and $I(\mu, 0) < +\infty$ if $\langle \mu, r\rangle < +\infty$, we get that $\mathcal{I}(\mu)$ is finite if and only if $\langle \mu, r\rangle$ is finite. Finally, note that since $I(\cdot, \cdot)$ is good, then the map $L^1_+(E) \ni Q \to I(\mu, Q)$

is lower semicontinuous with compact level sets and therefore it has a minimum. The uniqueness of the minimizer follows from the fact that $I(\mu, \cdot)$ is strictly convex on the set $\{Q : I(\mu, Q) < +\infty\}$, as can be easily derived from the strictly convexity of $\Phi(\cdot, p)$ for $p > 0$. The non trivial task is therefore to prove (1.13).

We will use the following characterization of I proved in [2] (see formula (5.4) there)

$$I(\mu, Q) = \sup I_{\phi,F}(\mu, Q)\,. \tag{1.14}$$

In (1.14) the supremum is among all pairs ϕ, F with $\phi \in L^\infty(V)$, $F \in L^\infty(E)$, being respectively $L^\infty(V)$ the set of bounded functions on vertices and $L^\infty(E)$ the set of bounded functions on edges. Moreover we have

$$I_{\phi,F}(\mu, Q) := \langle \phi, \operatorname{div} Q\rangle - \langle \mu, r^F - r\rangle + \sum_{(y,z)\in E} Q(y,z)F(y,z) \tag{1.15}$$

where $r^F : V \to (0, +\infty)$ is defined by

$$r^F(y) = \sum_{z\in V} r(y,z)e^{F(y,z)} \quad \text{and} \quad \langle \phi, \operatorname{div} Q\rangle = \sum_{y\in V} \phi(y) \operatorname{div} Q(y).$$

In [2] formula (1.14) is proved with a slightly different class of functions but the argument can be clearly adapted to the present setting. See also Section 4 for computations similar to (1.14).

In the reversible case we have the following additional result:

Proposition 1.1. *Assume the same setting of Theorem* 1.2. *Suppose that the invariant measure* π *is also reversible, i.e.* $\pi(y)r(y,z) = \pi(z)r(z,y)$ *for all* $y, z \in V$. *Suppose that* $\mu \in \mathcal{P}(V)$ *is such that* $\mathcal{I}(\mu) < +\infty$ *(i.e.* $\langle \mu, r\rangle < +\infty$*). Then*

$$Q^*(y,z) = Q^*(z,y) = \sqrt{\mu(y)\mu(z)r(y,z)r(z,y)}$$

is the minimizing flow in (1.12) *and it holds*

$$\mathcal{I}(\mu) = \frac{1}{2} \sum_{y\in V}\sum_{z\in V} \left(\sqrt{\mu(y)r(y,z)} - \sqrt{\mu(z)r(z,y)}\right)^2 .$$

Moreover (1.13) *admits a maximizing sequence* $g^{(n)}$ *that is a suitable approximating sequence in* $L^\infty(V)$ *of the extended function* $g : V \to \{-\infty\} \cup \mathbb{R}$ *defined by*

$$g(y) := \log \sqrt{\mu(y)/\pi(y)}\,. \tag{1.16}$$

The rest of the paper is devoted to the proof of Theorem 1.2 (see Section 2) and the proof of Proposition 1.1 (see Section 3). Most of the technical difficulties come from the case of V infinite. In Section 4 we give for V finite an alternative proof of Theorem 1.2 showing that it is indeed a special case of the Fenchel–Rockafellar Theorem. In the case $|V| < +\infty$ different proofs where given in [7] and [1].

2. Proof of Theorem 1.2

2.1. Some preliminary results on oriented graphs

Let $(\mathcal{V},\mathcal{E})$ be an oriented graph. Given $y,z\in\mathcal{V}$, an *oriented path* from y to z in $(\mathcal{V},\mathcal{E})$ is a finite string $(x_1,x_2,\dots,x_n)$ with $x_1=y$, $x_n=z$ and $(x_i,x_{i+1})\in\mathcal{E}$ for all $i=1,\dots,n-1$. A *cycle* in $(\mathcal{V},\mathcal{E})$ is an oriented path $(x_1,x_2,\dots,x_n)$ with $x_1=x_n$. It is called *self-avoiding* if $x_i\neq x_j$ for $1\le i<j<n$. Given a cycle C we denote by $1\!\!1_C$ the function on $\mathcal{E}$ taking value 1 on the edges (x_i,x_{i+1}), $1\le i<n$, and zero otherwise.

Let us now refer to the oriented graph (V,E). We denote by $\mathcal{C}$ the family of self-avoiding cycles in (V,E). Given $C\in\mathcal{C}$, note that $1\!\!1_C$ is a divergence-free flow in $L^1_+(E)$. In [2][Lemma 4.1] it is proved that any divergence-free flow $Q\in L^1_+(E)$ can be written as $Q=\sum_{C\in\mathcal{C}}\widehat{Q}(C)1\!\!1_C$ for suitable nonnegative constants $\widehat{Q}(C)$, $C\in\mathcal{C}$. The above decomposition has to be thought as $Q(y,z)=\sum_{C\in\mathcal{C}}\widehat{Q}(C)1\!\!1_C(y,z)$ for each edge $(y,z)\in E$.

Take $\mu\in\mathcal{P}(V)$ such that $\langle\mu,r\rangle<+\infty$. Consider the oriented graph (V_μ,E_μ) where

$$
\begin{aligned}
&E_\mu=\{(y,z)\,:\,Q^\mu(y,z)=\mu(y)r(y,z)>0\}\,,\\
&V_\mu=\{y\in V\,:\,\exists z\in V \text{ with } (y,z)\in E_\mu \text{ or } (z,y)\in E_\mu\}\,.
\end{aligned}
$$

Trivially, the support of μ is included in V_μ. If $z\in V_\mu\setminus\mathrm{supp}(\mu)$ then there exists $y\in\mathrm{supp}(\mu)$ with $r(y,z)>0$.

On the set V_μ we define the equivalence relation $y\sim z$ as follows: $y\sim z$ if and only if in (V_μ,E_μ) there exists an oriented path from y to z as well as an oriented path from z to y. We call $(V_\mu^{(\ell)})_{\ell\in\mathcal{L}}$ the equivalence classes of V_μ under the relation $\sim$, and set

$$
E_\mu^{(\ell)}:=\{(y,z)\in E_\mu\,:\,y,z\in V_\mu^{(\ell)}\}\,.
$$

Above $\mathcal{L}$ is the index set of the equivalence classes, given by $\mathcal{L}=\mathbb{N}=\{1,2,\dots\}$ if there are infinite classes, or $\mathcal{L}=\{1,2,\dots,|\mathcal{L}|\}$ if there is a finite number of classes.

Given $Q\in L^1_+(E)$ and given $y\neq z$ in V we set $Q(y,z)=0$ if $(y,z)\notin E$. The support of Q, denoted by $E(Q)$, is defined as $E(Q)=\{(y,z)\in E\,:\,Q(y,z)>0\}$.

Lemma 2.1. *Let $Q\in L^1_+(E)$ satisfy $I(\mu,Q)<+\infty$. Then $Q(y,z)=0$ if $y\not\sim z$ in (V_μ,E_μ). In particular, $E(Q)\subset\cup_{\ell\in\mathcal{L}}E_\mu^{(\ell)}$.*

Proof. Suppose that $Q(y,z)>0$. Since $I(\mu,Q)<+\infty$ the flow Q must be divergence-free. In addition, since $\Phi(q,p)=+\infty$ if $q>0$ and $p=0$, the flow Q must have support contained in E_μ. By the cyclic decomposition, we can write

$Q = \sum_{C \in \mathcal{C}} \widehat{Q}(C) \mathbb{I}_C$. In particular, $(y,z) \in C_0$ for some $C_0 \in \mathcal{C}$ with $\widehat{Q}(C_0) > 0$. Since Q has support contained in E_μ it must be $(u,v) \in E_\mu$ for all $(u,v) \in C_0$. The cycle C_0 can be divided in two oriented paths, one from y to z and one from z to y in (V_μ, E_μ). This implies that y, z belong to the same equivalence class $V_\mu^{(\ell)}$. Since $E(Q) \subset E_\mu$ it must be $(y,z) \in E_\mu$ and therefore $(y,z) \in E_\mu^{(\ell)}$. $\square$

2.2. Proof of Theorem 1.2

We define the functions $I_1, J_1 \colon \mathcal{P}(V) \to [0, +\infty]$ as the r.h.s. of (1.12) and (1.13), respectively:

$$I_1(\mu) := \inf \big\{ I(\mu, Q) \, : \, Q \in L^1_+(E) \big\} \,, \tag{2.1}$$

and

$$J_1(\mu) := \begin{cases} \sup \big\{ -\langle \mu, e^{-g} L e^g \rangle \, : \, g \in L^\infty(V) \big\} \,, & \text{if } \langle \mu, r \rangle < +\infty, \\ +\infty & \text{otherwise} \,. \end{cases} \tag{2.2}$$

As already explained we only need to prove the equality $I_1(\mu) = J_1(\mu)$. By Remark 1.1 we can restrict to probability measures μ such that $\langle \mu, r \rangle < +\infty$.

We first show the inequality $I_1(\mu) \geq J_1(\mu)$. Given $g : V \to \mathbb{R}$, we define the gradient $\nabla g : E \to \mathbb{R}$ as $\nabla g(y,z) := g(z) - g(y)$. We now observe that for any $g \in L^\infty(V)$ and for any divergence-free flow $Q \in L^1_+(E)$ the following integration by parts formula is satisfied:

$$\begin{aligned} \langle Q, \nabla g \rangle &= \sum_{(y,z) \in E} Q(y,z)(g(z) - g(y)) \\ &= \sum_{y \in V} g(y) \Big(\sum_{z:(z,y) \in E} Q(z,y) - \sum_{z:(y,z) \in E} Q(y,z) \Big) \\ &= -\langle \operatorname{div} Q, g \rangle = 0 \,. \end{aligned} \tag{2.3}$$

After these observations the conclusion is simple. We can restrict the infimum in (2.1) to divergence-free Q's. Fix $\phi, g \in L^\infty(V)$. We can use the variational characterization (1.14) of the rate function I and deduce

$$\begin{aligned} & I(\mu, Q) \geq I_{\phi, \nabla g}(\mu, Q) \\ & \quad = \sum_{(y,z) \in E} \big\{ -\mu(y) r(y,z) \big[e^{g(z) - g(y)} - 1 \big] + Q(y,z)[g(z) - g(y)] \big\} \\ & \quad = -\langle \mu, e^{-g} L e^g \rangle \,. \end{aligned} \tag{2.4}$$

In both the above identities we used that Q is divergence-free. Minimizing over Q and maximizing over g in (2.4) we obtain that $I_1(\mu) \geq J_1(\mu)$.

We show now the converse inequality $I_1(\mu) \leq J_1(\mu)$. As already observed after Theorem 1.2 the infimum defining I_1 in (2.1) is attained as some flow Q^*, i.e. $I(\mu, Q^*) = I_1(\mu)$. Taking $Q \equiv 0$ in (2.1) we get

$$I_1(\mu) = I(\mu, Q^*) \leq I(\mu, 0) = \langle \mu, r \rangle < +\infty \,. \tag{2.5}$$

Due to Lemma 2.1 we can write

$$I(\mu, Q^*) = \sum_{\ell \in \mathcal{L}} \sum_{(y,z) \in E_\mu^{(\ell)}} \Phi\Big(Q^*(y,z), \mu(y) r(y,z)\Big) + \sum_{(y,z) \in E_\mu \setminus \cup_{\ell \in \mathcal{L}} E_\mu^{(\ell)}} \mu(y) r(y,z). \tag{2.6}$$

The next step is to show that $Q^*(y,z) > 0$ for any $(y,z) \in E_\mu^{(\ell)}$. Suppose by contradiction that there exists an edge $(y,z) \in E_\mu^{(\ell)}$ such that $Q^*(y,z) = 0$. Take a cycle C contained in (V_μ^ℓ, E_μ^ℓ) and containing the edge (y,z) (it exists by definition of the equivalence relation $\sim$). For any $\alpha \geq 0$ consider the perturbed flow $Q^*_\alpha := Q^* + \alpha 1\!\!1_C$. Then $I(\mu, Q^*_\alpha) < +\infty$. Moreover, the map $\mathbb{R}_+ \ni \alpha \to I(\mu, Q^*_\alpha) \in [0, +\infty)$ is continuous and C^1 on $(0, +\infty)$. Its derivative on $(0, +\infty)$ is given by

$$\frac{d}{d\alpha}\big[I(\mu, Q^*_\alpha)\big] = \sum_{(v,w) \in C} \log \frac{Q^*(v,w) + \alpha}{\mu(v) r(v,w)} \,. \tag{2.7}$$

The above derivative becomes strictly negative for α small enough and this contradicts the fact that Q^* is a global minimizer.

Consider now an arbitrary cycle C contained in $(V_\mu^{(\ell)}, E_\mu^{(\ell)})$. We have just proved that Q^* is strictly positive on the edges of C. Hence, now we can conclude that the above function $\mathbb{R}_+ \ni \alpha \to I(\mu, Q^*_\alpha) \in [0, +\infty)$ is C^1 on all $\mathbb{R}_+$ (zero included) where its derivative is given by (2.7). Since Q^* is a global minimizer the value $\alpha = 0$ is a local minimum and consequently the value of the derivative in correspondence of $\alpha = 0$ must be zero. From (2.7) we get

$$\sum_{(v,w) \in C} \log \frac{Q^*(v,w)}{\mu(v) r(v,w)} = 0 \,. \tag{2.8}$$

The validity of (2.8) for any cycle C contained in $(V_\mu^{(\ell)}, E_\mu^{(\ell)})$ implies that there exists a function $g_\ell : V_\mu^{(\ell)} \to \mathbb{R}$ such that

$$\log \frac{Q^*(y,z)}{\mu(y) r(y,z)} = g_\ell(z) - g_\ell(y) \,, \qquad \forall (y,z) \in E_\mu^{(\ell)} \,. \tag{2.9}$$

The function g_ℓ is determined up to an arbitrary additive constant in the following way. Let y^* be an arbitrary fixed element of $V_\mu^{(\ell)}$ and set $g_\ell(y^*) := 0$. For any $z \in V_\mu^{(\ell)}$ consider an arbitrary oriented path $(z_1, \ldots, z_n)$ going from y^*

to z in (V_μ, E_μ) and define

$$g_\ell(z) := \sum_{i=1}^{n-1} \log \frac{Q^*(z_i, z_{i+1})}{\mu(z_i) r(z_i, z_{i+1})} \,. \tag{2.10}$$

If we prove that g_ℓ is well defined, i.e. that the above definition (2.10) does not depend on the chosen path, then it is immediate to check (2.9).

To show that the definition is well posed, fix an oriented path $(u_1, \dots, u_k)$ from z to y^* in (V_μ, E_μ) (it exists since $y^* \sim z$). Then $C = (z_1, \dots, z_n, u_2, \dots, u_k)$ is a cycle going through y^*, z. It is trivial to check that all points in C are $\sim$-equivalent to y^*, z, hence C is a cycle in $(V_\mu^{(\ell)}, E_\mu^{(\ell)})$. Applying (2.8) we get

$$\sum_{i=1}^{n-1} \log \frac{Q^*(z_i, z_{i+1})}{\mu(z_i) r(z_i, z_{i+1})} = -\sum_{j=1}^{k-1} \log \frac{Q^*(u_j, u_{j+1})}{\mu(u_j) r(u_j, u_{j+1})} \,. \tag{2.11}$$

This shows that the l.h.s. does not depend on the particular oriented path $(z_1, \dots, z_n)$ from y^* to z, since the r.h.s. is path-independent. Hence g_ℓ is well defined.

The function g_ℓ does not necessarily belong to $L^\infty(V_\mu^{(\ell)})$, nevertheless we can improve a result similar to (2.3):

Claim 2.1. *The series* $\sum_{(y,z)\in E_\mu^{(\ell)}} Q^*(y,z)(g_\ell(z) - g_\ell(y))$ *is absolutely convergent and moreover*

$$\sum_{(y,z)\in E_\mu^{(\ell)}} Q^*(y,z)(g_\ell(z) - g_\ell(y)) = 0 \,. \tag{2.12}$$

Since the series is absolutely convergent the l.h.s. of (2.12) does not depend on the order of summation and therefore is well defined.

Proof of the claim. By the triangle inequality we have

$$q\,|\log(q/p)| \le \Phi(q,p) + |q - p| \,, \qquad q, p > 0 \,. \tag{2.13}$$

Using the inequality (2.13) we have

$$\begin{aligned} &\sum_{(y,z)\in E_\mu^{(\ell)}} Q^*(y,z)\,|g_\ell(z) - g_\ell(y)| \\ &\quad\le \sum_{(y,z)\in E(Q^*)\subset E_\mu} Q^*(y,z) \left| \log \frac{Q^*(y,z)}{\mu(y) r(y,z)} \right| \\ &\quad\le I(\mu, Q^*) + \|Q^* - Q^\mu\|_1 < +\infty \,. \end{aligned} \tag{2.14}$$

Above we have used (2.5) and the fact that $Q^\mu \in L^1_+(E)$ since $\langle \mu, r\rangle < +\infty$. This proves that the series in the l.h.s. of (2.12) is absolutely convergent. Trivially, using the cyclic decomposition $Q^* = \sum_{C\in\mathcal{C}} \widehat{Q}(C)\mathbb{1}_C$, this is equivalent to the bound

$$\sum_{(y,z)\in E_\mu^{(\ell)}} \sum_{\substack{C\in\mathcal{C}:\\ C \text{ is inside } (V_\mu^{(\ell)},E_\mu^{(\ell)})}} \widehat{Q}(C)|g_\ell(z)-g_\ell(y)| < +\infty\,. \tag{2.15}$$

Due to (2.15) and the properties of absolutely convergent series (in particular, their invariance under permutation of the addenda) we get

$$\begin{aligned}&\sum_{\substack{C\in\mathcal{C}:\\ C \text{ is inside } (V_\mu^{(\ell)},E_\mu^{(\ell)})}} \Big(\sum_{(y,z)\in C} \widehat{Q}(C)(g_\ell(z)-g_\ell(y))\Big)\\ &\qquad = \sum_{(y,z)\in E_\mu^{(\ell)}} \Bigg(\sum_{\substack{C\in\mathcal{C}\,:\,(y,z)\in C\,,\\ C \text{ is inside } (V^{(\ell)},E^{(\ell)})}} \widehat{Q}(C)(g_\ell(z)-g_\ell(y))\Bigg)\,. \end{aligned}\tag{2.16}$$

On the other hand, the l.h.s. of (2.16) is trivially zero since each sum inside the brackets is zero. The r.h.s. of (2.16) is simply the l.h.s. of (2.12) (recall Lemma 2.1), thus concluding the proof of our claim. □

Using (2.9) and (2.12) we obtain

$$\sum_{(y,z)\in E_\mu^{(\ell)}} \Phi\big(Q^*(y,z),\mu(y)r(y,z)\big) = \sum_{(y,z)\in E_\mu^{(\ell)}} \mu(y)r(y,z)\big(1-e^{\nabla g_\ell(y,z)}\big)\,. \tag{2.17}$$

Recall that $\mathcal{L}$ is the index set of the equivalence classes for $\sim$ in (V_μ, E_μ). We then consider the oriented graph $(\mathcal{L},\mathcal{E})$ where the oriented edges are given by the pairs (ℓ,ℓ') for which there exists an edge $(y,z)\in E_\mu$ such that $y\in V_\mu^{(\ell)}$ and $z\in V_\mu^{(\ell')}$. Then the graph $(\mathcal{L},\mathcal{E})$ is an oriented acyclic graph, i.e. it contains no cycles.

We can now conclude the proof. First we consider the case when $|\mathcal{L}|<+\infty$. Then by Proposition 1.4.3 in [5] the finite acyclic oriented graph $(\mathcal{L},\mathcal{E})$ admits an acyclic ordering of the vertices. This means that there exists a bijection $\widehat{h}:\mathcal{L}\to\mathcal{L}$ such that $\widehat{h}(\ell)<\widehat{h}(\ell')$ for any $(\ell,\ell')\in\mathcal{E}$. Then we define $h(\ell):= |\mathcal{L}|-\widehat{h}(\ell)+1$ to get a bijection $h:\mathcal{L}\to\mathcal{L}$ such that $h(\ell)>h(\ell')$ for any $(\ell,\ell')\in\mathcal{E}$.

Consider the sequence of functions $g^{(n)}\in L^\infty(V)\,, n\in\mathbb{N}$ defined by

$$g^{(n)}(y) := \begin{cases} g_\ell^{(n)}(y)+h(\ell)n & \text{if } y\in V_\mu^{(\ell)} \text{ for some } \ell\in\mathcal{L}\,,\\ 0 & \text{otherwise}\,,\end{cases} \tag{2.18}$$

where

$$g_\ell^{(n)}(y) := \begin{cases} g_\ell(y) & \text{if } |g_\ell(y)| \leq \dfrac{n}{3}\,, \\ \dfrac{g_\ell(y)}{|g_\ell(y)|}\dfrac{n}{3} & \text{otherwise}\,. \end{cases} \tag{2.19}$$

We finally get

$$\begin{aligned} J_1(\mu) &\geq \lim_{n\to+\infty} -\big\langle \mu, \exp\{-g^{(n)}\} L \exp\{g^{(n)}\}\big\rangle \\ &= \lim_{n\to+\infty} \sum_{(y,z)\in E_\mu} \mu(y) r(y,z) \big(1-\exp\{\nabla g^{(n)}(y,z)\}\big) \\ &= \lim_{n\to+\infty} \Big[\sum_{\ell\in\mathcal{L}} \sum_{(y,z)\in E_\mu^{(\ell)}} \mu(y) r(y,z)\big(1-\exp\{\nabla g_\ell^{(n)}(y,z)\}\big) \\ &\quad + \sum_{\ell\neq\ell'} \sum_{\substack{(y,z)\in E_\mu: \\ y\in V_\mu^{(\ell)},\, z\in V_\mu^{(\ell')}}} \mu(y) r(y,z) \\ &\quad \times \big(1-\exp\{g_{\ell'}^{(n)}(y) - g_\ell^{(n)}(z) + [h(\ell')-h(\ell)]n\}\big)\Big]\,. \end{aligned} \tag{2.20}$$

The above limit can be computed applying the Dominated Convergence Theorem. To this aim we first observe that

$$\begin{cases} \big|\nabla g_\ell^{(n)}(y,z)\big| \ \leq\ \big|\nabla g_\ell(y,z)\big| & \text{if } (y,z)\in E_\mu^{(\ell)}, \\ \operatorname{sign}\{\nabla g_\ell^{(n)}(y,z)\} = \operatorname{sign}\{\nabla g_\ell(y,z)\} & \text{if } (y,z)\in E_\mu^{(\ell)}, \\ g_{\ell'}^{(n)}(y) - g_\ell^{(n)}(z) + [h(\ell')-h(\ell)]\,n \leq -\dfrac{1}{3}n & \text{if } (y,z)\in E_\mu, y\in V_\mu^{(\ell)}, z\in V_\mu^{(\ell')}. \end{cases} \tag{2.21}$$

Note that in the second case we have used that $(\ell,\ell')\in\mathcal{E}$ thus implying that $h(\ell')-h(\ell)\leq -1$. Note moreover that due to (2.21) we can write

$$\big|1-\exp\{\nabla g_\ell^{(n)}(y,z)\}\big| \leq 1+\exp\{\nabla g_\ell(y,z)\}\,, \qquad (y,z)\in E_\mu^{(\ell)}\,.$$

Since $\langle\mu, r\rangle < +\infty$ and due to (2.9) we are allowed to apply the the Dominated Convergence Theorem. As a consequence, we get

$$\begin{aligned} J_1(\mu) &\geq \ \text{r.h.s. of (2.20)} \\ &= \sum_{\ell\in\mathcal{L}} \sum_{(y,z)\in E_\mu^{(\ell)}} \mu(y) r(y,z)\,(1-\exp\{\nabla g_\ell(y,z)\}) \\ &\quad + \sum_{(y,z)\in E_\mu\setminus\cup_{\ell\in\mathcal{L}} E_\mu^{(\ell)}} \mu(y) r(y,z) \\ &= I(\mu, Q^*) = I_1(\mu)\,. \end{aligned} \tag{2.22}$$

Note that the second equality is a byproduct of (2.6) and (2.17). This ends the proof of $I_1(\mu) = J_1(\mu)$ when $|\mathcal{L}| < +\infty$. A special case with $|\mathcal{L}| < +\infty$ is when $\mathrm{supp}(\mu) = V$. In this case $|\mathcal{L}| = 1$ since $E_\mu = E$ and the Markov chain ξ is irreducible.

We can now treat the general case. Let μ be an arbitrary probability measure on V. We want to show that $J_1(\mu) \geq I_1(\mu)$.

We first observe that $J_1(\pi) = 0$, where we recall that π is the unique invariant measure of the Markov chain ξ. Indeed, since $1 - e^x \leq x$ for all $x \in \mathbb{R}$, we have for any $g \in L^\infty(V)$

$$\sum_{(y,z)\in E} \pi(y)r(y,z)(1-\exp\{\nabla g(y,z)\}) \leq \sum_{(y,z)\in E} \pi(y)r(y,z)\nabla g(y,z) = 0\,. \tag{2.23}$$

The last equality in (2.23) follows by (2.3) since Q^π is a divergence-free element of $L^1_+(E)$. Equation (2.23) gives $J_1(\pi) \leq 0$ and the converse inequality is obtained selecting in (2.2) a constant function g. This concludes the proof that $J_1(\pi) = 0$.

Since $\mathrm{supp}(\pi) = V$ for any $c \in (0,1)$ $\mathrm{supp}(c\mu + (1-c)\pi) = V$ so that from the result obtained in the case $|\mathcal{L}| < +\infty$ we know that

$$J_1(c\mu + (1-c)\pi) = I_1(c\mu + (1-c)\pi)\,. \tag{2.24}$$

Since J_1 is defined as a supremum of convex functions it is a convex function, hence $J_1(c\mu + (1-c)\pi) \leq cJ_1(\mu) + (1-c)J_1(\pi) = cJ_1(\mu)$. Invoking (2.24) we get

$$J_1(\mu) \geq \liminf_{c\to 1} \frac{I_1(c\mu + (1-c)\pi)}{c} \geq I_1(\mu)\,. \tag{2.25}$$

In the last inequality we have used the lower semicontinuity of I_1.

3. Proof of Proposition 1.1

We will use both the equivalent representations (1.12) and (1.13) of the rate function $\mathcal{I}$. We take $\mu \in \mathcal{P}(V)$ with $\langle \mu, r\rangle < +\infty$ and recall equation (1.16) that defines the extended function $g : V \to [-\infty, +\infty)$ as $g(y) := \log\sqrt{\mu(y)/\pi(y)}$, with the convention $\log 0 := -\infty$. We consider also the sequence of functions $g^{(n)} \in L^\infty(V)$ defined by

$$g^{(n)}(y) := \begin{cases} g(y) & \text{if } |g(y)| \leq n\,, \\ \dfrac{g(y)}{|g(y)|}n & \text{if } |g(y)| > n\,, \end{cases}$$

where by $(-\infty)/(+\infty)$ we mean -1. We observe that (use $2ab \leq a^2 + b^2$)

$$\sum_{y\in V}\sum_{z\in V} \mu(y)r(y,z)e^{\nabla g(y,z)} = \sum_y\sum_z \sqrt{\mu(y)r(y,z)\mu(z)r(z,y)} \leq \langle \mu, r\rangle < +\infty.$$

The above bound and the inequality

$$|1-\exp\{\nabla g^{(n)}(y,z)\}| \le 1+\exp\{\nabla g(y,z)\}$$

allows to apply the Dominated Convergence Theorem. As a consequence, by an elementary computation we get

$$\lim_{n\to+\infty} -\langle \mu, \exp\{-g^{(n)}\} L \exp\{g^{(n)}\}\rangle = \sum_{\{y,z\}\in E^u} \left(\sqrt{\mu(y)r(y,z)} - \sqrt{\mu(z)r(z,y)}\right)^2, \tag{3.1}$$

where with E^u we denote the set of unordered bonds. Since we are considering the reversible case then necessarily if $(y,z)\in E$ then also $(z,y)\in E$. Consequently $\{y,z\}\in E^u$ when both (y,z) and (z,y) belong to E. Equation (3.1) implies that $\mathcal{I}(\mu)$ is greater or equal to the right hand side of (3.1).

To prove the converse inequality we restrict the infimum in (1.12) to symmetric flows, i.e. to flows Q such that $Q(y,z)=Q(z,y)$ for any $(y,z)\in E$. Then we can define $S: E^u \to \mathbb{R}^+$ by $S(\{y,z\}) := Q(y,z) = Q(z,y)$. For symmetric flows the rate function $I(\mu,Q)$ can be written as

$$\begin{aligned} \sum_{\{y,z\}\in E^u} \Big[2S(\{y,z\}) \log \frac{S(\{y,z\})}{\sqrt{\mu(y)r(y,z)\mu(z)r(z,y)}} \\ + \mu(y)r(y,z) + \mu(z)r(z,y) - 2S(\{y,z\}) \Big]. \end{aligned} \tag{3.2}$$

The minimization procedure in (3.2) is easy since the zero divergence constraint is always satisfied. We can then solve an independent variational problem for each unordered bond, without any constraint apart the non-negativity of S. On the bond $\{y,z\}$ the minimizer is

$$S^*(\{y,z\}) = \sqrt{\mu(y)r(y,z)\mu(z)r(z,y)}\,.$$

Calling Q^* the associated symmetric flow we get that $I(\mu,Q^*)$ coincides with the right hand side of (3.1). This completes the proof.

4. Alternative proof of Theorem 1.2 for V finite

As explained after Theorem 1.2 we only need to show the identity $I_1 = J_1$ (recall (2.1) and (2.2)).

In the finite case an interesting proof of this result is obtained observing that it is a special case of the Fenchel – Rockafellar Theorem (see for example [4]). In the case $|V| = +\infty$ this strategy does not work since the continuity requirement in the following general statement is missing. Consider a topological vector space X and its dual X^*. Let $\phi, \psi : X \to (-\infty, +\infty]$ be two proper (i.e. not identically equal to $+\infty$) extended convex functions such that $\phi+\psi$ is proper and there exists an $x_0 \in X$ where either $\phi(x_0) < +\infty$ and ϕ is continuous at x_0 or $\psi(x_0) < +\infty$ and ψ is continuous at x_0. The Fenchel – Rockafellar Theorem states that

$$\inf_{x\in X} \{\phi(x) + \psi(x)\} = \sup_{f\in X^*} \{-\phi^*(-f) - \psi^*(f)\} \,, \tag{4.1}$$

where, given $\gamma : X \to (-\infty, +\infty]$,

$$\gamma^*(f) := \sup_{x\in X} \{\langle f, x\rangle - \gamma(x)\} \,, \qquad f \in X^* \,,$$

is the Legendre transform of γ.

Fix $\mu \in \mathcal{P}(V)$ and recall the definition of the graph (V_μ, E_μ) given at the beginning of Section 2, as well as the equivalence relation $\sim$ leading to the equivalence classes $V_\mu^{(\ell)}$, $\ell \in \mathcal{L}$, and associated set of edges $E_\mu^{(\ell)}$. Since V is finite the condition $\langle \mu, r\rangle < +\infty$ is automatically satisfied.

We want to apply the Fenchel – Rockafellar Theorem with $X = L^1(E_\mu)$ endowed of the standard L^1-norm and $X^* = L^\infty(E_\mu)$. Clearly in the finite case we could work with the simpler choice $X = X^* = \mathbb{R}^{E_\mu}$ with the Euclidean topology (we use instead a more general notation having in mind some possible extensions to the infinite case). Our choice for the functions ϕ, ψ is

$$\phi(Q) := \begin{cases} \displaystyle\sum_{(y,z)\in E_\mu} \Phi\big(Q(y,z), Q^\mu(y,z)\big) & \text{if } Q \in L^1_+(E_\mu)\,, \\ +\infty & \text{otherwise;} \end{cases}$$

$$\psi(Q) := \begin{cases} 0 & \text{if } \operatorname{div} Q = 0\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

Given $Q \in L^1(E_\mu)$ the divergence $\operatorname{div} Q : V \to \mathbb{R}$ is still defined as

$$\operatorname{div} Q(y) = \sum_{(y,z)\in E_\mu} Q(y,z) - \sum_{(z,y)\in E_\mu} Q(z,y)\,.$$

The above functions ϕ, ψ are proper convex functions (recall that $\Phi(\cdot, p)$ is convex for any $p \geq 0$). Moreover, since $\Phi(\cdot, p)$ is a continuous function on $(0, +\infty)$ for all $p > 0$, we conclude that $\phi(Q) < +\infty$ and ϕ is continuous at Q for any $Q \in L^1(E_\mu)$ such that $Q(y,z) > 0$ for all $(y,z) \in E_\mu$. Finally, we note that the function $\phi + \psi$ is proper since finite on the zero flow. Note that working with E instead of E_μ, neither ψ nor ϕ would have satisfied the condition of boundedness and continuity in at least one point.

Clearly it holds

$$I_1(\mu) = \inf_{Q\in L^1_+(E)} I(\mu, Q) = \inf_{Q\in L^1(E_\mu)} \{\phi(Q) + \psi(Q)\} \,. \tag{4.2}$$

The fact that $I_1(\mu) = J_1(\mu)$ then follows as a byproduct of the Fenchel–Rockafellar Theorem and the following Claims 4.1 and 4.2.

Claim 4.1. *For any $f \in L^\infty(E_\mu)$ we have*

$$\phi^*(f) = \sum_{(y,z)\in E_\mu} \mu(y)r(y,z)\big(e^{f(y,z)} - 1\big)\,, \qquad f \in L^\infty(E_\mu)\,. \tag{4.3}$$

Proof. For each edge $(y,z) \in E_\mu$, the function $\mathbb{R}_+ \ni u \to f(y,z)u - \Phi(u, Q^\mu(y,z))$ has maximum value $\mu(y)r(y,z)\left(e^{f(y,z)} - 1\right)$ attained at $u = e^{f(y,z)}Q^\mu(y,z) = e^{f(y,z)}\mu(y)r(y,z)$. □

Before stating Claim 4.2 we prove a geometric characterization of gradient functions on oriented graphs. To this aim we fix some language.

We define E^*_μ as the set of oriented edges (y,z) such that $(y,z) \in E_\mu$ or $(z,y) \in E_\mu$. Given a function f on E_μ we can extend it to a function f_* on E^*_μ setting

$$f_*(y,z) := \begin{cases} f(y,z) & \text{if } (y,z) \in E_\mu\,, \\ -f(z,y) & \text{if } (y,z) \in E^*_\mu \setminus E_\mu\,. \end{cases}$$

We say that $(x_1, x_2, \dots, x_n)$ is a generalized path from x_1 to x_n in (V_μ, E_μ) if for any $i = 1, \dots, n-1$ it holds $(x_i, x_{i+1}) \in E^*_\mu$ (in other words, $(x_1, x_2, \dots, x_n)$ is an oriented path in (V_μ, E^*_μ)). Given a generalized path γ we define $\int_\gamma f$ as

$$\int_\gamma f := \sum_{i=1}^{n-1} f_*(x_i, x_{i+1})$$

if γ is given by $x_1, x_2, \dots, x_n$.

Lemma 4.1. *Given $f \in L^\infty(E_\mu)$ there exists a $g : V \to \mathbb{R}$ such that $f = \nabla g$, i.e. such that $f(y,z) = g(z) - g(y)$ for all $(y,z) \in E_\mu$, if and only if for any pair of generalized paths γ and γ' having the same initial point and the same final point it holds*

$$\int_\gamma f = \int_{\gamma'} f\,. \tag{4.4}$$

Proof. If is simple to check that if $f = \nabla g$ then $f_*(y,z) = g(z) - g(y)$. This implies that $\int_\gamma f = g(z) - g(y)$ for any generalized path γ in (V_μ, E_μ) from y to z. Therefore (4.4) is satisfied whenever γ, γ' have the same extremes.

Suppose on the other hand that (4.4) is satisfied for any γ, γ' having the same extremes. We introduce on V_μ the equivalent relation $\sim^*$ saying that $y \sim^* z$ if there exists a generalized path from y to z in (V_μ, E_μ). It is simple to check that we have indeed an equivalence relation. For each equivalence class $W \subset V_\mu$ we fix a reference site $x_* \in W$ and define g on W setting $g(y) := \int_\gamma f$ where γ is any generalized path from x_* to y. Due to (4.4) the definition is well posed. Let us check that $\nabla g = f$. To this aim we fix $(y,z) \in E_\mu$. Clearly this implies $y \sim_* z$. Call x_* the reference site of their equivalence class and fix a path γ from x_* to y. If $\gamma = (x_1, x_2, \dots, x_n)$ set $\widetilde{\gamma} := (x_1, x_2, \dots, x_n, z)$. The path $\widetilde{\gamma}$ is a generalized path from x_* to z. By definition $g(z) = \int_{\widetilde{\gamma}} f = f(y,z) + \int_\gamma f = f(y,z) + g(y)$, thus concluding the proof. □

We can now state our final claim:

Claim 4.2. *For any $f \in L^\infty(E_\mu)$ we have*

$$\psi^*(f) = \begin{cases} 0 & \text{if } f = \nabla g \text{ for some } g : V \to \mathbb{R}\,, \\ +\infty & \text{otherwise}\,. \end{cases} \tag{4.5}$$

Proof. By a simple integration by parts it is trivial to check that it holds $\langle f, Q\rangle = 0$ if $f \in L^\infty(E_\mu)$ is of gradient type (i.e. $f = \nabla g$) and $Q \in L^1(E_\mu)$ is a divergence-free flow (see (2.3)). As a consequence we get

$$\psi^*(f) = \sup_{Q \in L^1(E_\mu)} \{\langle f, Q\rangle - \psi(Q)\} = 0\,,$$

for any f of gradient type.

Conversely suppose that f is not of gradient type. Then, by Lemma 4.1 there exist two generalized paths $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_m)$ in (V_μ, E_μ) such that $x_1 = y_1$, $x_n = y_m$ and

$$\sum_{i=1}^{n-1} f_*(x_i, x_{i+1}) - \sum_{j=1}^{m-1} f_*(y_j, y_{j+1}) > 0\,. \tag{4.6}$$

The given $\lambda > 0$ we define the divergence-free $Q_\lambda \in L^1(E_\mu)$ as

$$Q_\lambda(y,z) := \begin{cases} \lambda & \text{if } (y,z) = (x_i, x_{i+1}) \ i = 1, \dots, n-1\,, \\ -\lambda & \text{if } (y,z) = (x_{i+1}, x_i) \text{ and } (x_i, x_{i+1}) \notin E_\mu \ i = 1, \dots, n-1, \\ -\lambda & \text{if } (y,z) = (y_j, y_{j+1}) \ j = 1, \dots, m-1\,, \\ \lambda & \text{if } (y,z) = (y_{j+1}, y_j) \text{ and } (y_j, y_{j+1}) \notin E_\mu \ j = 1, \dots, m-1\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Then $\langle f, Q_\lambda\rangle$ equals λ times the r.h.s. of (4.6), thus implying that

$$\lim_{\lambda \to +\infty} \langle f, Q_\lambda\rangle = +\infty.$$

In particular, we obtain

$$\psi^*(f) \geq \lim_{\lambda \to +\infty} (\langle f, Q_\lambda\rangle - \psi(Q_\lambda)) = +\infty\,.$$

This ends the proof of our claim. □

References

[1] P. Baldi and M. Piccioni (1999) A representation formula for the large deviation rate function for the empirical law of a continuous time Markov chain. *Statist. Probab. Lett.* **41** (2), 107–115.

[2] L. Bertini, A. Faggionato and D. Gabrielli (2014) Large deviations of the empirical flow for continuous time Markov chains. To appear in *Ann. Inst. H. Poincaré – Probabilités et Statistiques.* http://arxiv.org/abs/1210.2004.

[3] L. Bertini, A. Faggionato and D. Gabrielli Flows, currents and cycles for Markov chains: large deviation asymptotics. Preprint arXiv:1408.5477.

[4] H. Brezis (2011) *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York.

[5] J. Bang-Jensen and G. Gutin (2001) *Digraphs. Theory, Algorithms and Applications.* Springer Monographs in Mathematics. Springer-Verlag London, Ltd., London.

[6] M.D. Donsker and S.R.S. Varadhan Asymptotic evaluation of certain Markov process expectations for large time. *Comm. Pure Appl. Math.* (I) **28**, 1–47 (1975); (II) **28**, 279–301 (1975); (III) **29**, 389–461 (1976); (IV) **36**, 183–212 (1983).

[7] G. Kesidis and J. Walrand (1993) Relative entropy between Markov transition rate matrices. *IEEE Trans. Inform. Theory* **39** (3), 1056–1057.

[8] R.E. Megginson (1998) *An Introduction to Banach Space Theory.* Springer-Verlag, New York.

[9] J.R. Norris (1999) *Markov Chains.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.